

Remarks

Claims 1-27 are pending in the application. Claims 1-27 are rejected. Claims 1, 3, 5, 17-18, 20-23, and 27 are amended. The Specification is amended. No new subject matter is added. The rejections are respectfully traversed.

The Specification is amended to overcome the Examiner's objections.

Claims 3, 5, 17, 18, 20, and 22-27 are rejected under 35 U.S.C. 112, second paragraph as being indefinite.

Claims 3, 5, 17, 18, 20, 22, 23 and 27 are amended to overcome the rejections.

Claim 1 is amended to more distinctly claim the invention.

Claims 1-3, 6-15, 19, 21, and 27 are rejected under 35 U.S.C. 102(e) as being anticipated by Abdel-Mottaleb (U.S. 6,424,789 -- "Abdel").

The invention is frame based. Each frame of an input video has an associated input frame play time, and the input video has a total input video play time that is a sum of the input frame play times of all of the input frames. *Each of the input frames of the input video is classified* according to a selected content characteristic. An output frame play time is *allocated to each* of the plurality of input frames based on the content characteristic of each of the input frames to generate a plurality of output frames.

Abdel measures differences between frames in a shot to classify the shot. The invention looks at only one frame at a time to classify the frame.

To those of ordinary skill in the art, it is well known that there is a huge difference between a 'shot' as described by Abdel, and a 'single frame' as claimed. Abdul correctly defines a "shot" (or scene) consisting of a **sequence of uninterrupted related frames,**" see Abstract. Applicants assert that a 'sequence' as described by Abdul can never be a single frame as claimed.

That is, a shot includes **all** frames from when a shutter of a camera opens, until the shutter closes. At typical frame rate of 30 or 60 frames per second, it would be impossible to perceive a shot as described by Abdel if it were to include only a single frame as claimed. Typically, perceivable 'shots' must last from at least several seconds to several minutes. This means that 'shots' typically include at least hundreds if not thousands of frames.

Now, Abdul determines a level of activity in a shot, and then performs the selected speed change special effect by adding frames or deleting frames to produce a modified shot, see Abstract. The invention operates in each frame as it comes by. In the most extreme case, and a most liberal interpretation of Abdel the shot must include at least a first frame and a second frame, see column 2,

45 The video processing device comprises an image processor
capable of identifying the at least one shot and determining
a first activity level within at least a first portion of the at
least one shot, wherein the first activity level indicates a rate
of change of video content between at least a first frame and
50 a second frame in the at least a first portion. The image
processor performs the selected speed change special effect
by one of adding frames and deleting frames in the at least
a first portion in response to the first activity level
determination, thereby producing a modified at least one
55 shot.

There is a serious problem with the Abdel system and method. Abdel can never operate in real time. This is obvious. Abdel needs to process a shot before a new shot can be created. Thus, if the shot is a minute long, the viewer sees nothing until at least a minute later, when the result of his processing becomes available.

In contrast, the invention processes the video and frame at the same time. As each input frame is classified, play time is allocated to each corresponding output frame, and the output can be viewed, as is, before any more frames need to be processed, as in Abdel.

Abdel is shot based. Abdel classifies shots, see col. 7, lines 33-38, below:

The shot frames are
examined to determine the relative degree of movement
between successive frames and/or across a longer series of 35
frames within the shot. The shot (or a sub-portion of the
shot) may then be classified according to the level of activity
in the shot (or sub-portion).

See also, Figures 3, 4A and 4B. Particularly, see element 320 of Figure 3, element 415 of Figure 4A, and element 460 of Figure 4B, each describing shot based classification.

In claims 2-3, the content characteristic is based on low-level features of each of the input frames. The low-level features are selected from a group consisting of motion vectors, color, texture, brightness, contrast, spectral parameters, local and global motion, activity, trajectory, speed, acceleration, object shape, object size, number of objects, number of faces, pitch, volume, tempo, and combinations thereof. The low level features for a frame are derived from each input frame for each corresponding output frame in the invention. Abdel measures frame differences in shots, which is not what is claimed.

In claim 6, the allocating of the play time is dynamically varied while processing the video. In claim 21, a total output video play time is allocated for an output video and the output frame play times are allocated so that a sum of the output frame play times of the plurality of output frames is equal to the total output video play time of the output video. As stated above, play time is *frame based* in the invention. Abdul increases the speed of a shot, see col. 8,

35 Once the relative amount of movement within the shot is
classified as high to low, the speed of the shot is increased
in relation to the original shot by saving sub-samples of
frames at a high or low rate. If the shot is classified as
low-level activity, the sub-sample occurs at a comparatively
low rate (process step 420). The low sub-sample rate results
40 in a relatively smaller number of frames being saved and a
relatively larger number of frames being dropped or deleted
from the original shot. For example, if the activity level is
low in a (or a sub-portion of the shot), image processor 120
may save one out of every three frames and drop the other
45 two frames.

Above, Abdel classifies the amount of movement in a shot. This is further described in col. 9:

minutes. The rate at which frames are dropped during the fast-forward speed change is then initially calculated by image processor 120 in order to comply with the five minute time constraint. Image processor 120 may then loop through the video clip one or more times to make the final modified video clip play in five minutes or less. Image processor 120 can apply different weighting factors to high-activity level and low-activity level portions of the video clip in order to meet with the specified time duration.

30

There, Abdel must "loop through the video clip one or more times" to meet the 5 minute constraint. This is a further indication that Abdel cannot be a real time process, and that Abdul is a multi-pass operation. In contrast, the invention is a single pass operation.

In claim 7, the allocated output frame play time of each of the output frames is determined by sampling the input frames. In claim 8, the sampling is a down-sampling of the input frames. In claim 9, the sampling is an up-sampling of the input frames. In claim 10, up-sampled output frames are interpolated from the input frames. In claim 11, the sampling is a combination of down-sampling and up-sampling of the input frames. The invention determines the allocated output frame play time for each output frame based on the classified content characteristic of each of the input frames. Abdel never determines play time. Abdel determines activity for an entire sequence of frames in a shot and adds or drops frames based on the activity. Abdel can never anticipate what is claimed.

In claim 12, the allocated output frame play time of each of the output frames is determined by an output frame rate of the output frame. In claim 13, the output frame rate is increased for selected input frames. In claim 14, the output frame rate is decreased for selected input frames. The invention determines output frame play

time for *each output frame*. Abdel never changes the frame rate for individual frames as claimed. Abdel determines whether the shot is to be played in fast forward or slow motion, and adds or drops frames of the shot to achieve a desired degree of motion, see col. 6 below:

20 Since the
video transfer rate of the image medium is constant, the
degree of motion in these shots can be slowed by the
duplication and addition of selected frames. Conversely, the
motion associated with these shots can be increased through
25 the deletion or removal of selected frames all in accordance
with the principles of the present invention.

Claimed is allocating a rate on a per frame basis. Abdel adds or drops frames to control a degree of motion in the shots, which has nothing to do with a rate of a frame.

In claim 15, the content characteristics of each of the plurality of input frames is measured to determine the classification. Each frame is measured to classify the frame. Abdel classifies a shot. These are totally different operations.

In claim 19, the input video is partitioned into a plurality of segments, and the input video is processed on a per segment basis. Therefore, different segments may be processed according to different selected content characteristics. The input video is still processed on a frame basis, which is never described by Abdel.

In claim 21, a total output video play time is allocated for an output video and the output frame play times are allocated so that a sum of the output frame play times of the plurality of output frames is equal to the total output video play time of the

output video. As described in col. 9, Abdel must "loop through the video clip one or more times" to meet the 5 minute constraint. Abdel is a multi-pass operation. In contrast, the invention is a single pass operation. Abdel cannot anticipate the invention.

Claims 4 and 5 are rejected under 35 U.S.C. 103(a) as being unpatentable over Abdel in view of Mehrotra, et al. (U.S. 6,665,423 – "Mehrotra").

Mehrotra generates descriptions of video *objects* using high-level and low-level time varying characteristics of the objects. The invention generates output frames having an output frame play time based on the selected content characteristic of each of the input frames, which can be high-level features of each of the input frames selected from a group consisting of genre, dramatic intensity, humor content, action level, beauty, lyricism, musical intensity, educational quality, profundity, nudity, linguistic class, and combinations thereof. Mehrotra never produces output frames. Mehrotra uses features to generate descriptions of video objects. The invention uses features to produce output frames. Mehrotra has nothing to do with what is claimed.

Claim 16 is rejected under 35 U.S.C. 103(a) as being unpatentable over Abdel in view of Park, et al. (U.S. 6,597,738 – "Park").

Park generates motion descriptors by measuring motion across an entire video or respective frames using various first order statistics, see col. 22, below.

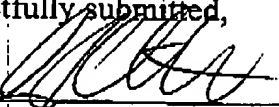
32 Next, the motion direction descriptor is to describe statistical characteristics of motion direction with relation to motion data for respective images or the whole images in the time which is intended by the time descriptor, and comprised of a motion direction average, central moment for the
40 average and standard deviation, dominant motion direction, accumulated motion histogram and direction data.

The invention determines higher order statistics for the measured characteristics of each frame to determine the classification for each frame to generate an output frame. Generating motion descriptors has nothing to do with generating an output frame having an output frame play time based on the selected content characteristic of an input frame as claimed. Park's motion descriptors can never be used to make the invention obvious.

All rejections have been complied with, and applicant respectfully submits that the application is now in condition for allowance. The applicant urges the Examiner to contact the applicant's attorney at phone and address indicated below if assistance is required to move the present application to allowance. Please charge any shortages in fees in connection with this filing to Deposit Account 50-0749.

Respectfully submitted,

Mitsubishi Electric
Research Laboratory, Inc.
201 Broadway
Cambridge MA, 02139
(617) 621-7539

By: 
Andrew J. Curtin
Reg. No. 48,485
Attorney for Assignee

ADAPTATION PROCESSING AUDIO BASED ON CONTENT
Title: CHARACTERISTICS OF PHONES IN THE AUDIO

Applicant: POWER et al Filed: 11/17/2004

Application No. 09/25,639 Patent No. _____ Docket No. MM-5001

Enclosed is the following:

☐ Transmittal Letter in triplicate;

☐ Application including Pgs. _____ of Spec. _____ Months;

☐ claims and Abstract;

☐ Declaration and Power of Attorney;

☐ Informal/Formal Drawings _____

☐ Sheets of Figs. _____

☐ Letter to Chief Draftsman;

☒ Amendment

☐ Information Disclosure Statement, Form 1449; _____ refs;

☐ Assignment, Recordation Cover Sheet;

Date Mailed: 7-19-2004

Date Due: 7-21-2004

Initials: PL

Express Mail Receipt No. _____

RECEIVED
JUL 21 2004
PTO Form 85B&C;
Year Maintenance Fee _____

Receipt Date & Application No. _____

U.S. PATENT & TRADEMARK OFFICE
JUL 21 2004

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.